

2.1 White space and Line Termination

In MiniC programs, the space, end-of-line, horizontal tab are known collectively whitespace characters. The end-of-line character marks the end of source program lines. A source line can be continued onto the next line by ending the first line with a backslash character. It is valid even in preprocessor command lines and within string constants. The splicing of source lines occurs before preprocessing and before lexical analysis of the MiniC program.

3. Comments

There are two ways to write a comment in MiniC. A comment begin with an occurrence of the two characters `/*` and ends with the first subsequent occurrence of the two characters `*/`. Comments are not recognized inside string or character constants or within other comments (i.e., cannot be nested). Comments may contain any number of characters.

The other type of comments begin with the two characters `//` and extends up to the next line break (but does not include it). All comments are to be replaced by a single space character for the purpose of further translation of MiniC program.

4. Tokens

Similar to C, there are five classes of tokens in MiniC: operators, separators, identifiers, keywords, and constants. The MiniC compiler forms the longest tokens possible when it collects characters in left-to-right order. Adjacent tokens may be separated by whitespace characters or comments.

4.1 Operators

Simple operators: `!` (not operator) `+` (Plus Sign) `*` (multiply) `-` (Minus)
`=` (Equal Sign) `|` (or operator) `<` (Less-than Sign) `>` (Greater-than Sign) `/`
(divide)

Compound assignment operators: `+=` `-=` `*=` `/=` `>=` `<=`

Other compound operators: `++` `--` `==` `!=`

4.2 Separators

Separator characters are: () [] , ; :

4.3 Identifiers

An identifier, or name, is a sequence of latin capital and small letters, digits, and underscore character. An identifier must not begin with a digit, and it must not have the same spelling as a keyword. Identifiers are case sensitive and the maximum length of an identifier is 255 characters.

4.4 Keywords

The following identifiers are keywords:

char, else, if, int, return, void, and while. They must not be used as ordinary identifiers.

4.5 Constants

There are three different kinds of constants in MiniC: integers, characters, and strings.